

Performance improvement of an optical network providing services based on multicast

Vincent Reinhard^{1,2} Johanne Cohen², Joanna Tomasik¹,
Dominique Barth², Marc-Antoine Weisser¹

(1) SUPELEC Systems Sciences, Computer Science Dpt.,
91192 Gif sur Yvette, France
email : FirstName.LastName@supelec.fr

(2) PRiSM, University of Versailles,
45 avenue des Etats-Unis, 78035 Versailles, France
email : FirstName.LastName@prism.uvsq.fr

January 20, 2013

Abstract

Operators of networks covering large areas are confronted with demands from some of their customers who are virtual service providers. These providers may call for the connectivity service which fulfils the specificity of their services, for instance a multicast transition with allocated bandwidth. On the other hand, network operators want to make profit by trading the connectivity service of requested quality to their customers and to limit their infrastructure investments (or do not invest anything at all).

We focus on circuit switching optical networks and work on repetitive multicast demands whose source and destinations are *à priori* known by an operator. He may therefore have corresponding trees “ready to be allocated” and adapt his network infrastructure according to these recurrent transmissions. This adjustment consists in setting available branching routers in the selected nodes of a predefined tree. The branching nodes are opto-electronic nodes which are able to duplicate data and re-transmit it in several directions. These nodes are, however, more expensive and more energy consuming than transparent ones.

In this paper we are interested in the choice of nodes of a multicast tree where the limited number of branching routers should be located in order to minimize the amount of required bandwidth. After formally stating the problem we solve it by proposing a polynomial algorithm whose optimality we prove. We perform exhaustive computations to show an operator gain obtained by using our algorithm. These computations are

made for different methods of the multicast tree construction. We conclude by giving dimensioning guidelines and outline our further work.

1 Introduction

Optical networks have become a dominant technology in modern networks covering large areas. Their advantage consists in providing an ultra-high bit rate obtained with slight energy consumption. All-optical networks are particularly interesting from economic and ecological point of view because a cost of transparent routers is low and their energy consumption is negligible [6].

Modern networks face a growing demand on the part of service providers. New offered services are more complex than the simple connectivity service assured traditionally by network operators. Providers sell services like storage and computation together with connectivity service to their customers. The part of this market ensuring on-the-fly resource allocation, called for commercial reasons Cloud Computing [1], is under a rapid development. In order to meet the demands of their customers, virtual service providers have to purchase a guaranteed connectivity service at network operators. At the same time, network operators can deal with numerous virtual service providers. They are interested in using their network resources the most efficiently and in this way minimize the cost of a prospective extension of their existing infrastructure.

We studied the mechanisms to execute distributed applications in an optical mesh network in the context of the CARRIOCAS project [2, 23]. Unlike a customary approach applied in Grids where applications benefit from a dedicated network infrastructure [9], this project went into the study of the coexistence of massive distributed applications in a network whose operator should make financial profit. With GMPLS [17] deployed, the CARRIOCAS network has to ensure both unicast and multicast transmissions. Routers which are able to duplicate data and send it in several directions allow a network operator to lower the bandwidth amount necessary to construct a multicast tree. On the other hand, these branching nodes are more expensive and more energy consuming than the transparent ones. The realistic assumption is thus that only a subset of routers is equipped with the duplicating functionality. In [19] we presented our solution to the problem consisting in the construction of a tree to any multicast request with minimization of the amount of used bandwidth under assumption of a limited number of branching nodes. The solution is heuristic because we proved that this problem is *NP*-complete. It turned out to be the most effective when the branching nodes were placed in the most homogeneous way in a network. The most homogeneous placement of k branching nodes represents in fact a solution to the k -centre problem which is also *NP*-complete [10].

Our study mentioned above inspired us to explore certain special cases of multicast demands. A network operator can know in advance recurrent multicast transmissions which require a lot of bandwidth. Being aware of frequent demands for identical (or almost identical) multicast transmissions an operator may have corresponding trees “ready to be allocated” and adapt his network infrastructure according to these recurrent transmissions.

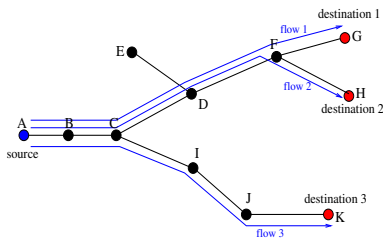


Figure 1: A multicast with source A and destinations G, H, K built up as a set of unicasts (without branching nodes)

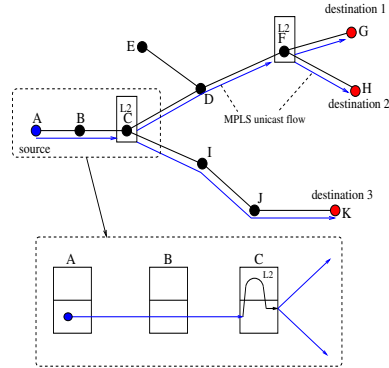


Figure 2: A multicast with source A , destinations G, H, K and branching nodes C, F

This adjustment may consist in setting available branching routers in the selected nodes of the predefined tree. In this paper we are interested in the choice of nodes of a multicast tree where the branching routers should be located in order to minimize the amount of required bandwidth. This approach allows an operator to make his network more efficient without any additional cost.

In the following section we make a survey of existing solutions to multicast tree allocation and explain the specificity of branching routers. In Section 3 our problem is stated in the formal way. We also formulate (Section 4) the solution properties. Next, we propose an algorithm to solve our problem, compute its complexity, and prove that it gives an optimal solution. Our problem is evidenced to be polynomial. Section 6 presents the results of bandwidth requirements for multicast trees depending on the number of available branching routers. The multicast trees which are subject of this analysis have been obtained by two methods, the first one based on the shortest path approach and the second one based on the Steiner tree approach. In the final section we give the conclusions and outline our further work.

2 Multicast tree construction

There are several schemes for multicasting data in networks [21, 12]. We present here the schemes adapted to optical circuit switching networks. The first one is to construct virtual circuits from the multicast source to each destination. Such a scheme is equivalent to multiple unicasts (Fig. 1) and the network bandwidth used by a large multicast group may become unacceptable [16].

In another scheme the multicast source sends data to the first destination and each destination acts as a source for the next destination until all destinations receive the data flow. In yet another scheme, intermediate routers make copies of data packets and dispatch them to their successors in the multicast tree. This solution allows the multicast transmission

to share bandwidth on the common links. Numerous multicast tree algorithms, which follow the latter scheme, have been proposed and can roughly be classified into two categories [21]. The first category contains the algorithms based on the shortest path while minimizing the weight of the path from the multicast source to each destination. The second category contains algorithms based on the Steiner tree problem [3, 5, 11, 14] which we formally define in Section 3. The algorithms derived from the Steiner tree problem minimize the total weight of the multicast tree. They are heuristic because the Steiner tree problem is NP-complete [11].

From the technological point of view, routers able to duplicate packets introduce a supplementary delay due to O/E/O conversions and are more expensive. For these reasons network operators want to limit the number of such routers which we call "diffusing nodes" or "branching nodes". The diffusing nodes which we consider are not equipped with the functionality "drop-and-continue" [25] as this operation mode is nowadays applied in practice exclusively in border routers. In Fig. 2 we go back to the example illustrated in Fig. 1. This time there are two branching nodes which allow one to reduce the amount of used bandwidth. Contrary to the solution built up of unicasts, in the one with branching nodes the bandwidth is used only once in each link.

3 Formalization of optimization problem

An optical network is modelled by a directed connected symmetrical graph [4] $G = (V, E)$. A multicast request is a pair $\epsilon = (e, R)$, where $e \in V$ is a multicast source and $R \subset V$ is a set of multicast destinations. We suppose that all multicast requests which we deal with can be transmitted in the network as a set of unicast transmissions (see Section 2). Therefore, we do not have to make precise the amount of data to transfer. For a given multicast request ϵ we first determine its tree, $A_\epsilon = (V_{A_\epsilon}, E_{A_\epsilon})$. This tree is a subgraph of G rooted in e , whose leaves are in the set R and whose arcs are directed from the root towards the leaves. We note D_{A_ϵ} the diffusing nodes in A_ϵ , $D_{A_\epsilon} \subseteq V_{A_\epsilon}$. Their allowed number is written as k . We now try to determine the choice of diffusing nodes in order to minimize the bandwidth consumption.

We will adopt as a metric of the bandwidth used by a multicast request a total number of arcs which construct its tree taking into account the fact that an arc may transport the same data more than once. To define this metric formally we start by determining the situations in which a request ϵ is satisfied by a set of paths in the multicast tree, $\mathcal{S}(D_{A_\epsilon})$. These situations are as follows:

- every node of R is the final extremity of exactly one path in $\mathcal{S}(D_{A_\epsilon})$,
- every node of D_{A_ϵ} is the final extremity of at most one path in $\mathcal{S}(D_{A_\epsilon})$,
- the origin of a path in $\mathcal{S}(D_{A_\epsilon})$ is either e or a node of D_{A_ϵ} ; in the latter case, the node of D_{A_ϵ} is also the final extremity of a path in $\mathcal{S}(D_{A_\epsilon})$,

- any node of $a \in D_{A_\epsilon}$ is in a path $p \in \mathcal{S}(D_{A_\epsilon})$ only if it is the final extremity or the origin of p .

The metric load_{A_ϵ} is defined as a sum of lengths of all paths in $\mathcal{S}(D_{A_\epsilon})$. The optimization problem which consists in placing k diffusing nodes can be thus formalized as:

theorem 1 Diffusing Nodes in Multicast Tree Problem (DNMTP)

Data: a directed connected symmetrical graph $G = (V, E)$, a multicast request $\epsilon = (e, R)$, a rooted multicast tree corresponding to this request A_ϵ , and a natural number k .

Goal: Find D_{A_ϵ} , $|D_{A_\epsilon}| \leq k$ so that load_{A_ϵ} is minimal.

4 Properties of the solution induced by the subset of vertices D

We now focus on a given multicast ϵ and we omit the subscript ϵ in the formulæ for their clarity. This section is devoted to studying properties of the solution induced by a set D of diffusing nodes, $\mathcal{S}(D)$. We introduce the notation used for its description. For any u , $u \in V_A$ in A we define A^u as a sub-tree of A rooted in u . We also define three parameters of u in A . A set D^u is a set of diffusing nodes in tree A^u ($D^u \subseteq D$). A set R^u is a set of destinations nodes in tree A^u ($R^u \subseteq R$). a^u is the arc connecting A^u from the remainder of A . We propose:

theorem 2 Let D be a set of vertices in A . Let u be a vertex in A . The path number $pn(u)$ is a number of paths in a solution $\mathcal{S}(D)$ spanned on A which pass through u or which terminate in u . The window of the solution $\mathcal{S}(D)$ on arc a^u is an triplet of integers (β, d, load) where β is its path number $pn(u)$, $d = |D^u|$ and load represents the load of $\mathcal{S}(D)$ in tree A^u .

We can notice that each solution induced by the set of diffusing nodes D , can be defined by each window for each arc of tree A .

theorem 1 Let D be a set of vertices in A . Let u be a node having one child u_1 of A . The window on arc a^u is equal to

$$\begin{cases} (1, d + 1, \text{load} + 1) & \text{if } u \in D \\ (b + 1, d, \text{load} + b + 1) & \text{if } u \notin D \text{ and } u \in R \\ (b, d, \text{load} + b) & \text{if } u \notin D \text{ and } u \notin R \end{cases} \quad (1)$$

where the window on arc a^{u_1} is (b, d, load) .

Proof: First, we assume that $u \in D$ (u is a diffusing node). By definition of path number, arc a^u has a path number equal to one. Since the window on arc a^{u_1} is (b, d, load) , tree A^u contains one more diffusing node than tree A^{u_1} .

Second, we assume that $u \notin D$. So, u is not a diffusing node and tree A^u contains exactly the same set of diffusing nodes as tree A^{u_1} . If $u \in R$, then u is the final extremity of exactly

one path. So the path number on arc a^u is equal to the path number on arc a^{u_1} plus one. If $u \notin R$, then the path number on arc a^u is equal to the path number on arc a^{u_1} .

From these statements, we can compute the load of the solution \mathcal{S} induced by D in tree A^u . The load increases by the path number on arc a^u .

□

Now, using the same arguments in the proof of Lemma 1, we extend it when u has several children.

theorem 2 *Let u be a node of A having ℓ children $u_1 \dots, u_\ell$. The window on arc a^u is equal to*

$$\begin{cases} (1, & 1 + \sum_i^\ell d_i, & \sum_i^\ell \text{load}_i + 1) & \text{if } u \in D \\ (1 + \sum_i^\ell b_i, & \sum_i^\ell d_i, & \sum_i^\ell \text{load}_i + b + 1) & \text{if } u \notin D \text{ and } u \in R \\ (\sum_i^\ell b_i, & \sum_i^\ell d_i, & \sum_i^\ell \text{load}_i + b) & \text{if } u \notin D \text{ and } u \notin R \end{cases} \quad (2)$$

where the window on arc a^{u_i} is $(b_i, d_i, \text{load}_i)$ for any i , $1 \leq i \leq \ell$.

Now, we want to compare two solutions by introducing a partial order for each node.

theorem 3 *Let D and D' be two subsets of vertices in A . Let v be a vertex. $\mathcal{S}(D) \preceq_v \mathcal{S}(D')$ if and only if the following three conditions are simultaneously satisfied: (i) $b \leq b'$; (ii) $d \leq d'$; (iii) $\text{load} \leq \text{load}'$, where the window of the solution $\mathcal{S}(D)$ (respectively $\mathcal{S}(D')$) on arc a^v is (b, d, load) (respectively (b', d', load')).*

theorem 1 *Let u be a vertex in A . Let D and D' be two subsets of vertices in A such that $\mathcal{S}(D) \preceq_u \mathcal{S}(D')$. Then the solution induced by D'' , where $D'' = (D' \setminus D^u) \cup D^u$, satisfies the following property*

$$\mathcal{S}(D'') \preceq_v \mathcal{S}(D') \text{ for all nodes } v \text{ not in } A^u$$

Proof: Let P be the path between root e and vertex v .

First, we focus on vertices v outside A^u and not in P . Since $D''^v = D'^v$, arc a^v has the same window of the solution induced by D and of the solution $\mathcal{S}(D')$.

Second, we focus on vertices v in P . By definition of the partial order \preceq_u , we have (i) $b \leq b'$, (ii) $d \leq d'$, and (iii) $\text{load} \leq \text{load}'$, where the window of the solution $\mathcal{S}(D)$ (respectively $\mathcal{S}(D')$) on arc a^u is (b, d, load) (respectively (b', d', load')). Now, we can compute the window of the solution $\mathcal{S}(D'')$ on arc a^t where t is the father of node u . Let $(b'_t, d'_t, \text{load}'_t)$ be the window of the solution $\mathcal{S}(D')$ on arc a^t .

From Lemma 2, if $t \in D'$, then the window of the solution $\mathcal{S}(D'')$ on arc a^t is $(1, d'_t - d' + d, \text{load}'_t - \text{load}' + \text{load})$. Thus $\mathcal{S}(D'') \preceq_t \mathcal{S}(D')$. We can apply the same arguments as previously for the other case. The same reasoning goes for each vertex of this path starting from the father of t until the root. This completes the proof of Property 1. □

theorem 4 *Let u be a vertex in A . Let D be a subset of vertices in A . D is sub-optimal for A^u if and only if for any D' which is a subset of vertices in A such that $d = d'$ and $b = b'$, we have $\text{load} \leq \text{load}'$ where the window of the solution $\mathcal{S}(D)$ (respectively $\mathcal{S}(D')$) on arc a^u is (b, d, load) (respectively (b', d', load')).*

theorem 2 *Let u be a vertex of A having ℓ children $u_1 \dots, u_\ell$. Let D be a subset of vertices in A . If D is sub-optimal for node u , then D is also sub-optimal for node u_i , for any integer i , $1 \leq i \leq \ell$.*

Proof: We can prove this property by contradiction. Assume that there is at least one child u_i of u such that D is not sub-optimal for node u_i . So it implies that there exists a subset D' such that D' is sub-optimal for node u_i and such that $d = d'$ and $b = b'$, we have $load' < load$ where the window of the solution $\mathcal{S}(D)$ (respectively $\mathcal{S}(D')$) on arc a^u is $(b, d, load)$ (respectively $(b', d', load')$). So, using Lemma 2, we can construct a subset D'' such that $D'' = (D \setminus D^{u_i}) \cup D'^{u_i}$ and such that $\mathcal{S}(D'') \preceq_u \mathcal{S}(D)$. So, it implies that D is not optimal. So there is a contradiction. \square

5 Algorithm, its complexity and optimality

Our algorithm is based on the dynamic approach. We introduce the notation used for its description. For any u , $u \in V_{A_\epsilon}$ in A_ϵ we define A_ϵ^u as a sub-tree of A_ϵ rooted in u . We also define two parameters of u in A_ϵ . The height $h(u)$ is a distance between u and e in A_ϵ . We also note $h_{\max} = \max_{u \in V_{A_\epsilon}} h(u)$. The path number $pn(u)$ is a number of paths in a solution $\mathcal{S}(D_{A_\epsilon})$ with a given set of diffusing nodes spanned on A_ϵ which pass through u or which terminate in u . It is obvious that if u is a branching node then $pn(u) = 1$.

The idea of our algorithm is to compute for any u , $u \in V(A)$, some sub-optimal sets D of diffusing nodes for A^u where the window of the solution $\mathcal{S}(D)$ on arc a^u is $(b, d, load)$. One set D is constructed for any value b , $1 \leq b \leq |R^u|$, any value d , $0 \leq d \leq k$. As the reader might already guess, a sub-optimal set D for the root e gives a solution to our problem. We want therefore to find these sets starting from the leaves and ending up in the root of A . As u may be or may not be a diffusing node, we have to know how to compute the two sets for both the cases.

As u may not be equipped with the branching property, the minimal load of the sub-optimal set for it should be stored in the matrix $M(u)$ whose rows are indexed by $pn(u)$ (these indices are $1, 2, \dots, |R|$) and whose columns are indexed by the number of diffusing nodes deployed in A^u (these indices are $0, 1, \dots, k$). If a solution does not exist, the corresponding matrix element is equal to zero.

As u may become a branching node, the minimal load of the sub-optimal set can be stored in a line vector $L(u)$ because the path number of a diffusing node is always equal to one.

In a nutshell: $M_{i,j}(u) = \alpha \neq 0$ ($L_i(u) = \alpha \neq 0$, respectively) if and only if a sub-optimal set D exists in A^u having its window on arc a^u equal to (j, i, α) . (respectively to $(1, i, \alpha)$). For computational reasons the destinations u , which are leaves of A , have “unitary” matrix and vector attributed: $M_{1,0}(u) = 1$, $L_1(u) = 1$ and all other elements are zero.

As we have said above, our algorithm to solve the DNMTTP attributes to each node u its $M(u)$ and $L(u)$ starting from the leaves whose height is $H = h_{\max}$ and performing the

Procedure Mat_Vec_Filling

1. **If** u is a leaf **then** attribute the “unitary” $M(u)$ and $L(u)$ to u **endIf**
2. **If** u is not a leaf **then**
3. choose arbitrarily v which is one of the successors of u in A_ϵ^u ;
4. **First_Succ_Mat_Vec**(u, v); mark v ;
5. **While** there is a successor of u in A_ϵ^u which has not be marked yet **do**
6. choose arbitrarily w among the non-marked successors of u in A_ϵ^u ;
7. **Others_Succ_Mat_Vec**(u, w); mark w
8. **endWhile**
9. **endIf**

Figure 3: The procedure **Mat_Vec_Filling**

bottom-up scheme with $H = H - 1$ until the root is reached ($H = 0$). The attribution of $M(u)$ and $L(u)$ to u is realised by the procedure **Mat_Vec_Filling** (Fig. 3). This procedure takes a node u and its corresponding sub-tree as data. Intuitively speaking, this is a modified breadth-first search [13] in which one arbitrarily chosen successor, treated first, computes its matrix and vector (the **First_Succ_Mat_Vec** procedure) in a different way from its brothers (the **Others_Succ_Mat_Vec** procedure). The leaves have the “unitary” matrix and vector assigned.

The procedure **First_Succ_Mat_Vec** operates on a node u and one of its successors v for which $M(v)$ and $L(v)$ are already known as **Mat_Vec_Filling** follows a bottom-up approach (Fig. 4). It uses the variable **elt** to store the non-zero elements in a column i of $M(u)$ and $L(u)$. The procedure executes the function \min^+ whose two arguments are natural. It returns a minimum of these two values in exception of the case in which one of the arguments is zero. The other positive argument is when returned. The main idea is based on the observation that the weight of the multicast tree in $A^v \cup \{u\}$ is equal to the multicast weight in A^v incremented by the weight of reaching u which is itself equal to $\text{pn}(u)$. Let us remind the reader that $\text{pn}(u) = 1$ when u is a diffusing node and $\text{pn}(u)$ is a matrix row index otherwise.

Remark 1: From Lemma 1 and Property 1, we can deduce, that if u has one child v for any $i, j, 0 \leq i \leq k$ and $1 \leq j \leq |R|$

- $L_i(u) = 1 + \min^+(L_{i-1}(v), \min^+\{M_{i-1,j}(v) : j : 1 \leq j \leq |R|\})$
- $M_{i,1}(u) = 1 + \min^+(L_i(v), M_{i,1}(v))$
- $M_{i,j}(u) = j + M_{i,j'}(v)$ where $j \neq 1$, and $j' = j - 1$ if $u \in R$, otherwise $j' = j$

The procedure **First_Succ_Mat_Vec** computes the formulæ here above.


```

Procedure First_Succ_Mat_Vec(u, v)
1.  $L_1(u) = \min_j^+(M_{0,j}(v))$ 
2. ForAll  $i$  such that  $0 < i \leq k$  do
3.    $L_i(u) = 1 + \min_j^+(\min_j^+(M_{i-1,j}(v), L_{i-1}(v))$ 
4. endForAll
5. ForAll  $i$  such that  $0 \leq i \leq k$  do
6.   ForAll  $j$  such that  $0 < i \leq |R|$  do
7.     If  $j == 1$  then  $\text{elT} = 1 + \min^+(M_{i,1}(v), L_i(v))$ 
8.     else  $\text{elT} = j + M_{i,j}(v)$ 
9.     endIf;
10.    If  $u$  is destination of multicast  $\epsilon$  then
11.       $M_{i,j+1}(u) = \text{elT} + 1$  else  $M_{i,j}(u) = \text{elT}$ 
12.    endIf
13. endForAll
14. endForAll

```

Figure 4: The procedure **First_Succ_Mat_Vec**

In lines 1–4 $L(u)$ is computed for u seen as a diffusing node. On the i^{th} step the smallest positive weight is chosen between weights of its predecessor v seen as a diffusing and a non-diffusing node. These weights are taken for v with one diffusing node less because u itself is diffusing. This weight is increased by the weight of reaching u which is one as u is diffusing. Lines 5–9 fill up $M(u)$ when u is seen as non-diffusing. Line 7 treats the case in which only one path passes through or terminates in u . The successor of u can be either a diffusing or non-diffusing node. Otherwise (line 8) its successor has to be a non-diffusing node. The case in which u is a destination despite the fact that it is not a leaf in A is treated in lines 10–12 as the weight of the access to u has to be added.

Others_Succ_Mat_Vec (Fig. 5) operates on a node u and its successors w different from v which has already been examined in **First_Succ_Mat_Vec**. The procedure uses the variable **elT** as **First_Succ_Mat_Vec** does. Furthermore, the procedure makes use of the auxiliary variables $M'(u)$ and $L'(u)$ to store the new values of $M(u)$ and $L(u)$ as the current elements of $M(u)$ and $L(u)$ are still in use. The procedure **Others_Succ_Mat_Vec** is built up on the same principle as the previous one. Lines 1–9 treat the filling up of $L(u)$ and lines 10–23 treat the filling up of $M(u)$. The important difference consists in traversing all the couples (x, y) , $x, y = 1, 2, \dots, k$ or $x, y = 0, 1, \dots, k$ such that $x + y = i$. It leads from the fact that this time the weight of the multicast tree in $A^v \cup A^w \cup \{u\}$ is equal to the sum of the multicast weights in $A^v \cup \{u\}$ and in A^w with the branching nodes deployed in both $A^v \cup \{u\}$ and A^w . The matrix computation also requires an appropriate path number in order to determine the additional tree weight (lines 15–19).

Procedure Others_Succ_Mat_Vec(u,w)

1. **ForAll** i such that $0 < i \leq k$ **do**
2. $\text{elT} = \infty$;
3. **ForAll** (x, y) such that $(x, y) \in \{1, 2, \dots, k\} \times \{1, 2, \dots, k\}$ and $x + y = i$ **do**
4. **If** $(L_x(u) + L_y(w)) < \text{elT}$ **then** $\text{elT} = L_x(u) + L_y(w)$ **endIf**;
5. **If** $(L_x(u) + \min_j^+(M_{y,j}(w))) < \text{elT}$ **then** $\text{elT} = L_x(u) + \min_j^+(M_{y,j}(w))$
6. **endIf**;
7. **endForAll**
8. $V'_i(u) = \text{elT}$
9. **endForAll**;
10. **ForAll** i such that $0 \leq i \leq k$ **do**
11. $\text{elT} = \infty$;
12. **ForAll** (x, y) such that $(x, y) \in \{0, 1, \dots, k\} \times \{0, 1, \dots, k\}$ and $x + y = i$ **do**
13. **ForAll** j such that $0 < j \leq |R|$ **do**
14. $\text{elT} = \infty$;
15. **ForAll** (a, b) such that $(a, b) \in \{0, 1, \dots, k\} \times \{0, 1, \dots, k\}$ and $a + b = j$ **do**
16. **If** $M_{x,a}(u) + M_{y,b}(u) + b < \text{elT}$ **then**
17. $\text{elT} = M_{x,a}(u) + M_{y,b}(u) + b$
18. **endIf**
19. **endForAll**;
20. $M'_{i,j}(u) = \min^+(\text{elT}, M_{x,j-1}(u) + L_y(w) + 1)$
21. **endForAll**
22. **endForAll**
23. **endForAll**
24. $L(u) \leftarrow L'(u); M(u) \leftarrow M'(u);$

Figure 5: The procedure Others_Succ_Mat_Vec

Remark 2: From Lemma 2 and Property 2, we can deduce, that if u has ℓ children u_1, \dots, u_ℓ for any i, j , $0 \leq i \leq k$ and $1 \leq j \leq |R|$

- $L_i^\ell(u) = \min^+ \{L_{i'}^{\ell-1}(u) + L_{i''}(u_\ell), L_{i'}^{\ell-1}(u) + \min_{j: 1 \leq j \leq |R|}^+ M_{i'',j}(v) : i' + i'' = i\}$
- $M_{i,j}^\ell(u) = \min^+ \left(\begin{array}{ll} \min^+ \{M_{i',j'}^{\ell-1}(u) + M_{i'',j''}(u_\ell) + j'' & : i' + i'' = i \wedge j' + j'' = j\}, \\ \min^+ \{M_{i',j-1}^{\ell-1}(u) + L_{i''}(u_\ell) + 1 & : i' + i'' = i\}, \end{array} \right)$

where $M^f(u)$ and $L^f(u)$ correspond to the matrix and the vector computed by the algorithm for the sub-tree of A^u where u has only f children u_1, \dots, u_f .

The procedure `Others_Succ_Mat_Vec` computes the formulæ here above.

theorem 1 *The optimal set of diffusing nodes is obtained by the configuration associated to $\min_{i: 1 \leq i \leq k} L_i(e)$. Its complexity is $\mathcal{O}(k^2 |R|^2 |V_A|)$.*

Proof: From Remarks 1 and 2, we can deduce from any u in $V(A)$, the algorithm `Mat_Vec_Mat_Filling` computes vector $L(u)$ and matrix $M(u)$ such that $\forall b, 1 \leq b \leq |R|$, $\forall d, 0 \leq d \leq k$, thus there are two sub-optimal sets of diffusing nodes: one has load $L_d(u)$ and the other has load $M_{d,b}(u)$. \square

6 Numerical results

Our algorithm determines the optimal localizations for k diffusing nodes in a multicast tree which has already been created for a request $\epsilon = (e, R)$. As we have signalled in Section 2 there are numerous methods of construction of these trees. We selected two heuristic methods in order to observe their impact on the efficiency of our algorithm. The first one establishes a shortest path (ShP) between e and each $r \in R$. The corresponding multicast tree A_ϵ^{ShP} is a union of these shortest paths. The second method, which is based on the 2-approximable solution of the Steiner tree problem proposed in [22], gives A_ϵ^{StT} tree. This Steiner problem formalized in terms of multicast demand can be written as:

theorem 5 Steiner Tree Problem (StTP)

Data: a connected undirected graph $G = (V, E)$, a multicast request $\epsilon = (e, R)$, and a natural number k .

Question: Does a rooted tree A_ϵ exist such that the number of its arc is less than or equal to k ?

The heuristic algorithm [22] leans on polynomial algorithms of a minimum-weight spanning tree [15] and of a shortest path [8] coupled.

To generate a graph of 200 nodes we apply the Waxman model [24] of BRITE [18] (with default parameters). We estimate with the 5% precision at the significance level $\alpha = 0.05$ the average weight of multicast tree as a function of the destination number for both the algorithms which construct a tree. For each number of destinations we choose uniformly in

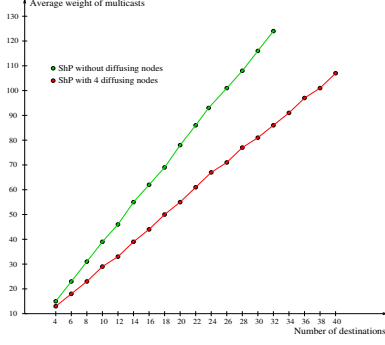


Figure 6: Average $A_{\epsilon}^{\text{ShP}}$ weight as a function of the number of destinations with and without diffusing nodes

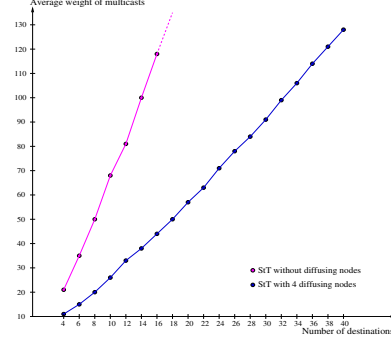


Figure 7: Average $A_{\epsilon}^{\text{StT}}$ weight as a function of the number of destinations with and without diffusing nodes

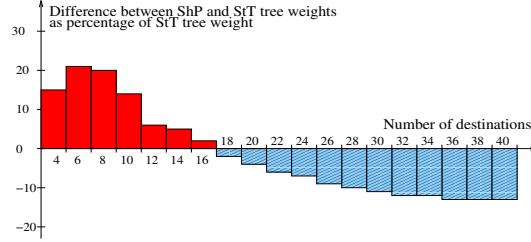


Figure 8: Difference between weights of $A_{\epsilon}^{\text{ShP}}$ and $A_{\epsilon}^{\text{StT}}$ as $A_{\epsilon}^{\text{StT}}$ weight percentage with four diffusing nodes as a function of the number of destinations

V a multicast source e and next, we select the destinations of this source according to the uniform distribution in $V - \{e\}$.

In order to perceive the impact of diffusing nodes on the tree weight we perform the computations with four nodes placed by our algorithm, and without them. In Fig. 6 we observe that the weight reduction obtained for ShP with the diffusing nodes is significant (about 31% for 32 destinations). The improvement obtained by the introduction of diffusing nodes into the trees built with StT (Fig. 7) is even more substantial than in the previous case (about 65% for 16 destinations). These two figures exhibit that ShP generates trees whose weight is less than those generated by StT. This fact is not astonishing as ShP always chooses a shortest path between the source and any destination.

In Fig. 8, in which the relative difference between $A_{\epsilon}^{\text{ShP}}$ and $A_{\epsilon}^{\text{StT}}$ weights as StT tree weight percentage is depicted, we notice, however, that this tendency is inverse for multicast trees with few destination (up to 18). To explain this phenomenon we notice that 1) with a small number of destinations the shortest paths identified by ShP are disjoint, and 2) typically, the edges of a tree obtained by shortest paths are more numerous than those of a Steiner tree computed for an identical multicast demand.

We now fix the number of destinations to 20 and we estimate the weights of trees obtained with ShP and StT algorithms in function of the number of branching nodes. We remind

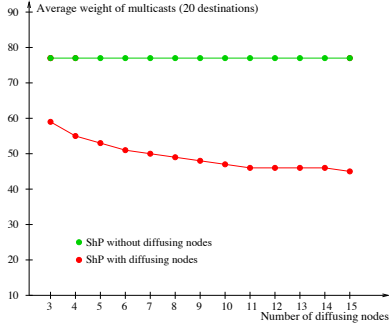


Figure 9: Average $A_{\epsilon}^{\text{ShP}}$ weight as a function of the number of diffusing nodes for 20 multicast destinations

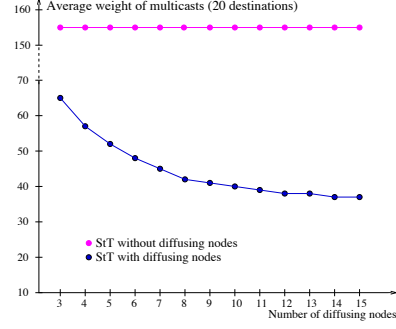


Figure 10: Average $A_{\epsilon}^{\text{StT}}$ weight as a function of the number of diffusing nodes for 20 multicast destinations

the reader that for 20 destinations and 4 diffusing nodes ShP turned out to be slightly more efficient than StT. Figs. 9 and 10 also show the average weight of ShP and StT trees estimated with the absence of diffusing nodes. In accordance with the comment made above in the context of the absence of diffusing nodes, ShP trees are almost twice as good as StT ones.

The introduction of three diffusing nodes reduces the weight of $A_{\epsilon}^{\text{ShP}}$ by about 20% (Fig. 9). Further additions allows one to lower the tree weight by almost 40% for 15 branching nodes. The influence of the branching nodes on the reduction of the tree weight in the StT case is striking (Fig. 10): an improvement of almost 60% in the case of three branching nodes until almost 75% for 15 of them. Confronting the results of ShP and StT with diffusing nodes we observe that StT, despite its starting point at a worse position, reaches the tree weight of 40 in the situation in which ShP has this weight of 48.

In Fig. 11 we observe the relative difference of ShP and StT tree weights for 20 multicast destinations in function of the number of diffusing nodes. It is not surprising that for this relatively large number of destinations and few diffusing nodes StT exhibits better performance than ShP. When the number of branching nodes increases and approaches the number of destinations, ShP trees become lighter than StT ones for the same reasons as those mentioned in the comments on Fig. 8.

The next question we ask ourselves concerns the detection of the numbers of diffusing nodes and destinations up to which StT is more advantageous than ShP. For the network investigated above the critical point is (4, 18). In Fig. 12 we mark critical points starting from which the ShP tree gives “lighter” solutions. For the points above the line we recommend ShP (for example, for three branching nodes and 30 destinations), for those below the line we recommend StT.

As the critical points depicted in Fig. 12 form a straight line whose slope is five, we are now interested in what this gradient depends on. One may guess that it is determined by the average degree of the network. Indeed, if we look at Fig. 13, the gradient decreases as the average node degree increases. Consequently, the line seen in Fig. 12 inclines with the average degree growth. Therefore we conclude that StT is more favourable for loosely

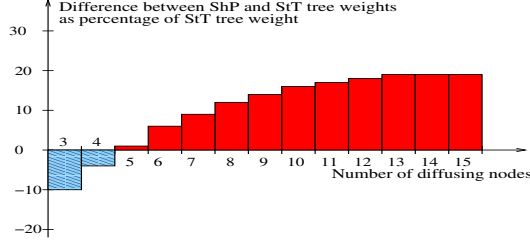


Figure 11: Difference between weights of ShP and StT trees as a percentage of StT tree weight as a function of the number of diffusing nodes for 20 multicast destinations

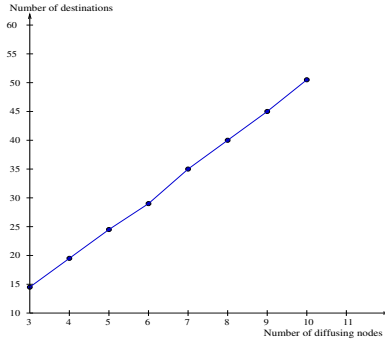


Figure 12: Critical point line determining the utility of ShP and StT trees

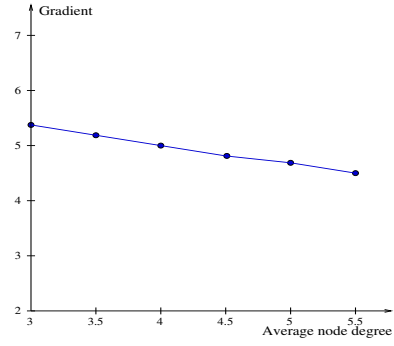


Figure 13: Gradient of critical point line in function of the average graph degree

connected graphs and ShP is better for dense networks.

7 Conclusions and further work

We studied a problem of infrastructural design of a commercial optical meshed network with a circuit switching routing protocol deployed. This problem was stated within the context of virtual services based on multicast transmission. It concerns frequent and voluminous multicast demands whose source and destinations are *à priori* known and its solution determines the localizations of branching nodes (i.e. routers with higher cost and energy consumption but which allow one to duplicate data and retransmit it in different directions). A solution to this problem allows a network operator to use his available resources more efficiently and make more profit with less, or even without any, investment.

After formally stating the problem we proposed an algorithm to solve it. Next, we proved its optimality and computed its complexity which is polynomial. We computed a gain in terms of the used bandwidth compared with multicast trees without any diffusing nodes. Among the two heuristic algorithms which we used to deploy multicast trees the first is based on the shortest path approach (ShP) and the second one exploits a solution to the Steiner tree problem (StT) in undirected graphs. We performed exhaustive computations

in order to compare the efficiency of our algorithm for multicast trees built with ShP and StT. We observed the dependency of their efficiency on the numbers of diffusing nodes and destinations. This dependency is influenced by the average network degree. StT works better in loosely connected networks whereas ShP is more efficient for strongly connected ones.

Generally speaking, we found ShP more efficient in finding a multicast tree than StT. We should not forget, however, that we used the 2-approximable algorithm. It is not excluded that a more precise StT algorithm (for example [7, 20]) may give better results. We consider implementing these algorithms in order to verify their performance for our purposes.

We plan to continue this work in order to determine a specific solution in particular graphs (for example having bounded treewidth). We conjecture that our algorithm could be extended to this kind of graph. On the other hand we consider pursuing our work on optimal multicast deployment by studying the Steiner problem in certain oriented graphs.

References

- [1] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia. A view of Cloud Computing. *Comm. ACM*, 53:50–58, April 2010.
- [2] O. Audouin. CARRIOCAS description and how it will require changes in the network to support Grids. In *20th Open Grid Forum*, 2007.
- [3] J. Beasley. An SST-based algorithm for the Steiner problem in graphs. *Networks*, 19, 1989.
- [4] C. Berge. *The theory of graphs and its applications*. Wiley, 1966.
- [5] K. Bharath-Kumar and J. M. Jaffe. Routing to multiple destinations in computer networks. *IEEE Trans. Commun.*, 31:343–351, March 1983.
- [6] E. Bonetto, L. Chiaraviglio, D. Cuda, G. Gavilanes Castillo, and F. Neri. Optical technologies can improve the energy efficiency of networks. In *ECOC*, 2009.
- [7] J. Byrka, F. Grandoni, T. Rothvoß, and L. Sanita. An improved LP-based approximation for Steiner tree. In *ACM-STOC*, 2010.
- [8] E. W. Dijkstra. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1:269–271, 1959.
- [9] I. T. Foster, C. Kesselman, J. M. Nick, and S. Tuecke. Grid services for distributed system integration. *IEEE Computer*, 35(6):37–46, 2002.
- [10] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman and company, 25th edition, 1979.

- [11] F. K. Hwang and D. S. Richards. Steiner tree problems. *Networks*, 22(1), 1992.
- [12] M. Jeong, C. Qiao, Y. Xiong, H. C. Cankaya, and M. Vandenhoude. Efficient multicast schemes for optical burst-switched WDM networks. In *ICC*, 2000.
- [13] D. E. Knuth. *The Art Of Computer Programming, vol. 1*. Addison-Wesley, 1997.
- [14] V. P. Kompella, J. Pasquale, and G. C. Polyzos. Multicasting for multimedia applications. In *INFOCOM*, 1992.
- [15] J. B. Kruskal. On the shortest spanning subtree of a graph and the traveling salesman problem. *Proc. of AMS*, 7(1):48–50, February 1956.
- [16] R. Malli, X. Zhang, and C. Qiao. Benefit of multicasting in all-optical networks. In *SPIE All Optical Networking*, pages 209–220, 1998.
- [17] E. Mannie. RFC 3945 — GMPLS, October 2004.
- [18] A. Medina, A. Lakhina, I. Matta, and J. Byers. BRITE: An approach to universal topology generation. In *MASCOTS*, Cincinnati, OH, USA, August 2001.
- [19] V. Reinhard, J. Tomasik, D. Barth, and M-A. Weisser. Bandwidth optimisation for multicast transmissions in virtual circuits networks. In *IFIP Networking*, 2009.
- [20] G. Robins and A. Zelikovsky. Improved Steiner tree approximation in graphs. In *ACM-SIAM*, 2000.
- [21] H. F. Salama, D. S. Reeves, and Y. Viniotis. Evaluation of multicast routing algorithms for real-time communication on high-speed networks. *IEEE J. on Sel. Areas in Comm.*, 15(3):332–345, 1997.
- [22] H. Takahashi and A. Matsuyama. An approximate solution for the Steiner problem in graphs. *Math. Jap.*, 24(6):573–577, 1980.
- [23] D. Verchère, O. Audouin, B. Berde, A. Chiosi, R. Douville, H. Pouylau, P. Primet, M. Pasin, S. Soudan, D. Barth, C. Caderé, V. Reinhard, and J. Tomasik. Automatic network services aligned with grid application requirements in CARRIOCAS project. In *GridNets*, pages 196–205, 2008.
- [24] B. M. Waxman. Routing of multipoint connections. *J-SAC*, 6(9), 1988.
- [25] X. Zhang, J. Y. Wei, and C. Qiao. Constrained multicast routing in WDM networks with sparse light splitting. *J. of Lightwave Tech.*, 18(12):1917–1927, 2000.